TEAM

PHOENIX

CONTENTS

1. INTRODUCTION

We are given the dataset for the thrust-force and torque signals captured by a dynamometer during the drilling cycle for various drill bits. Tests were conducted on HAAS VF-1 CNC Machining Center with Kistler 9257B piezo-dynamometer (sampled at 250Hz) to drill holes in ¼ inch stainless steel bars. Each drill-bit was used until it reached a state of physical failure.

The problem statement is to build a prediction model to predict the RUL (Remaining Useful Life) during the machining process in real life. RUL can be understood as the

2. DATA OVERVIEW

We have been provided with the data of 14 drill bits with varying number of holes drilled. For example, drill 1 has 21 holes whereas drill 14 has 45 holes. Each combination of drill bit and hole has been represented uniquely. For example, d1h1 corresponds to hole 1 of drill 1 and so on. Each dataset of these holes of respective drill bits have two columns. While one column has the data of thrust force during the process of drilling, the other column has the data of torque signals as the drilling progressed. Notably, these two variables namely 'thrust force' and 'torques signal' have been measured with the passage of time. So, they are basically function of time (as the drilling process goes on)

We have divided this dataset as first 10 drill bits and last 4 drill bits for training and testing respectively.

3. SUMMARY OF THE GIVEN RAW DATA

- Total number of drill bits = 14
- Total number of individual data sheets = 333

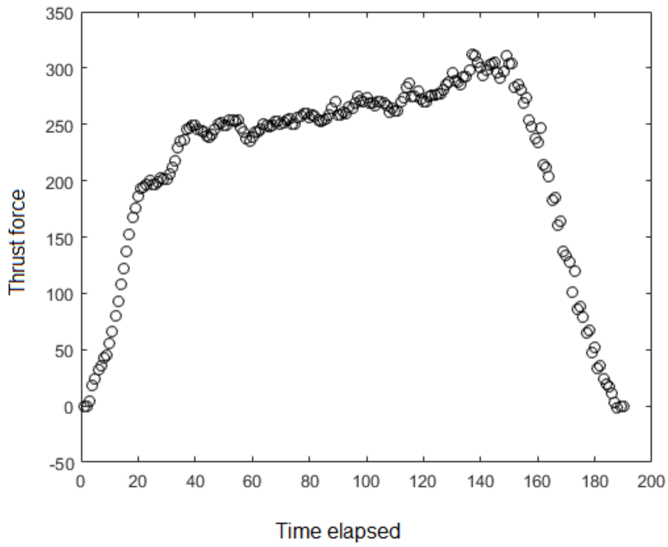
1. The sensor signals are in the form of time series segments.

2. The data sheet of respective holes of drill bits contained these data:
    - Thrust force
    - Torque signal

3. The length of the observation sequence i.e., the range of time over which data has been recorded is not different for different drill bits and holes.

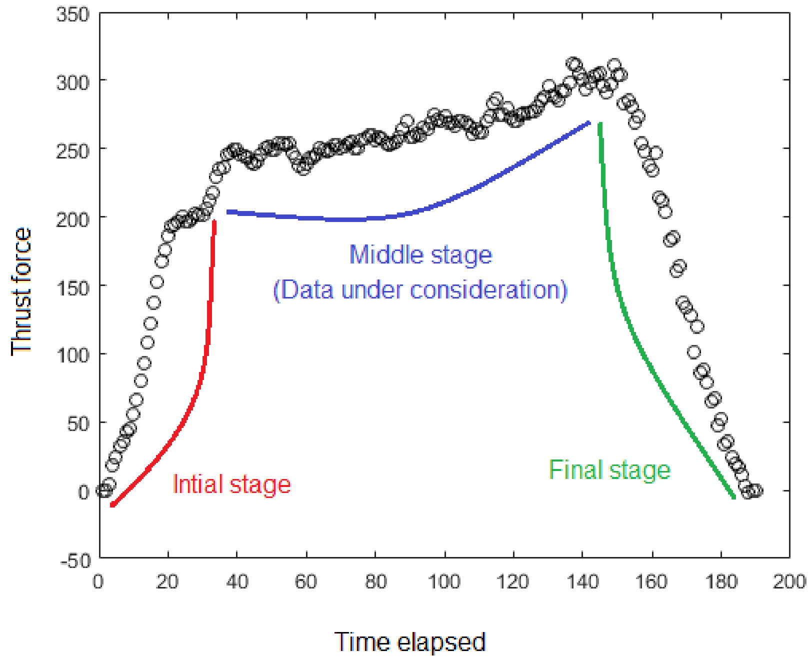
4. DATA PRE-PROCESSING

- Our aim is to extract the data which is useful for shaping the prediction model out of the entire dataset. The general plot of 'thrust force' versus 'time' for d1h1 is:

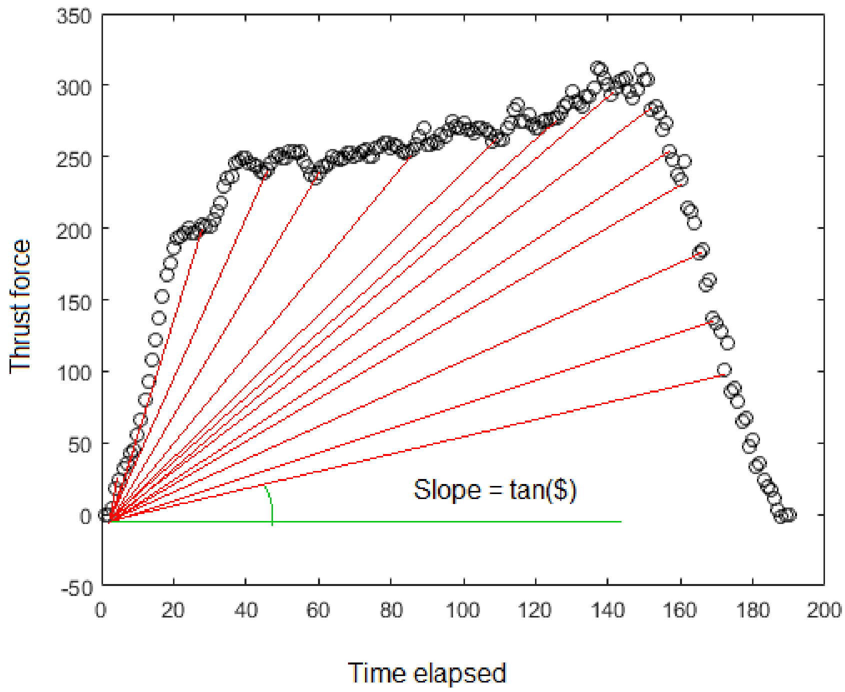


It can be observed in the above graph that the variation of the thrust force with time can be analyzed in three parts. First, the value of thrust force increases rapidly initially. After that, it follows a very less fluctuating approximate straight line. And at last, it drops down to its minimum value very rapidly. We intend to take those data which form that approximate straight line as the entire process of drilling that hole mostly depends on those data. **The reason behind this is that the actual drilling process occurs in the middle stages where maximum thrust force and torque are required and that's where the drill bit is most likely to fail.** The data at other both ends of this middle segment represents initialization of drilling and finishing of drilling.

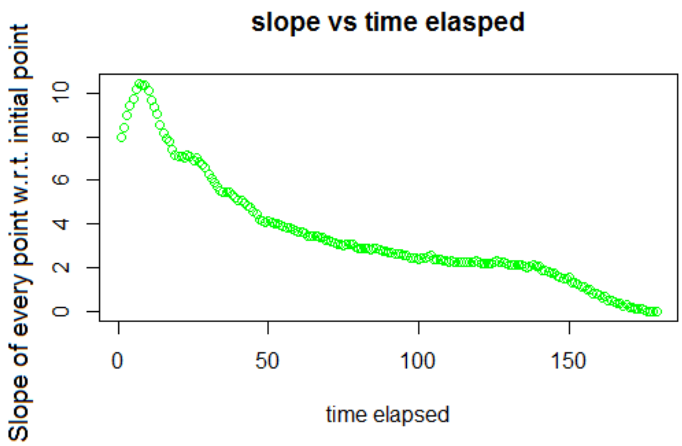
- To extract the desired data, we intend to remove the data at initial stages where it's increasing very rapidly as well as the data at final stages. After removing these two sets of data, we will be left with the desired set of data on the basis of which we can create our prediction model.



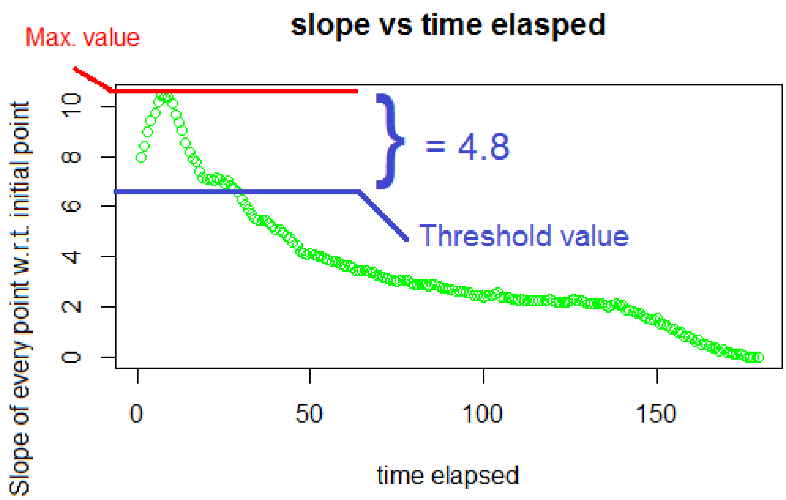
➤ Now, we calculate the slope of each line joining the other data points to the first data point.

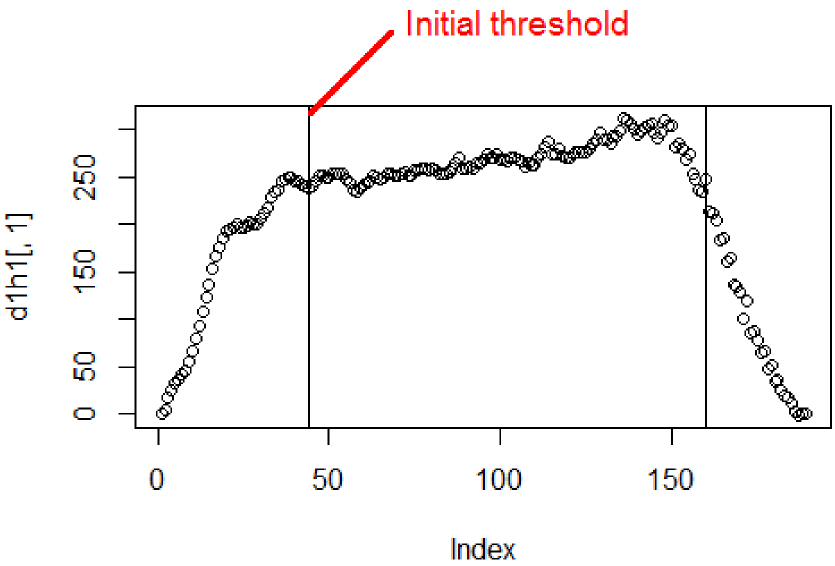


➤ The graph of slope versus time is as shown below:

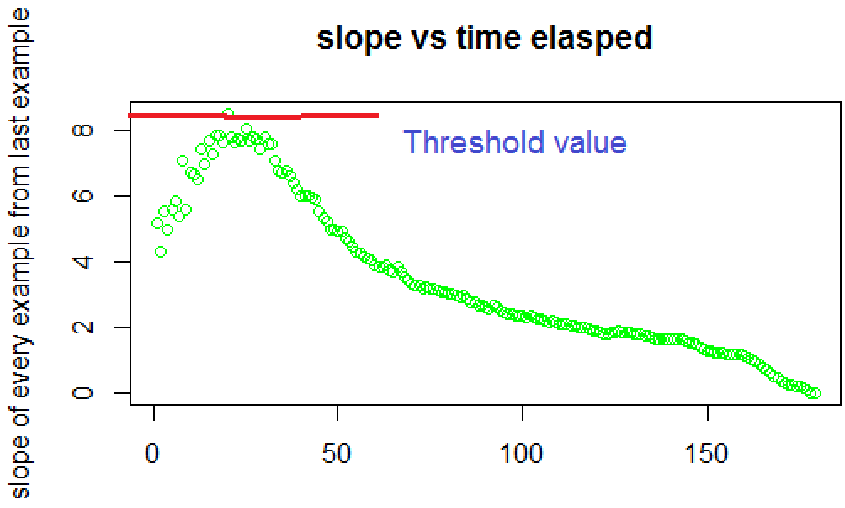


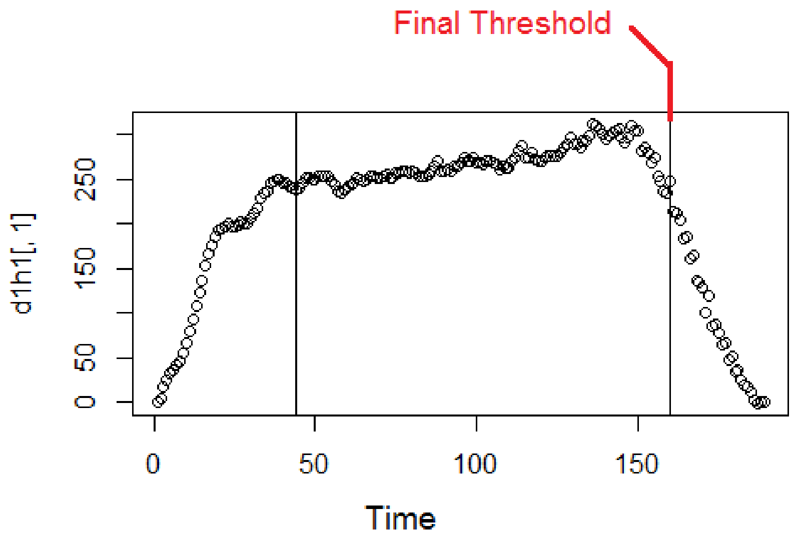
➤ After trying out different combination of values, we found that if we take a threshold around 4.8-5.0 below the maximum value then we get **a proper initial threshold** for the original data i.e., the data after this initial threshold will be considered to lie in middle stage region.



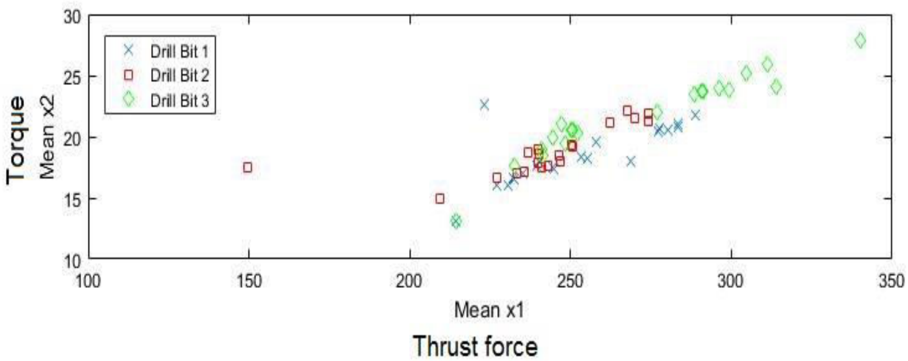


- The same thing we did to get final threshold value. We took slope of every point with respect to the last (final) point. We set the maximum value of the slop as threshold itself. And corresponding to that point, we get our final threshold point in original dataset.



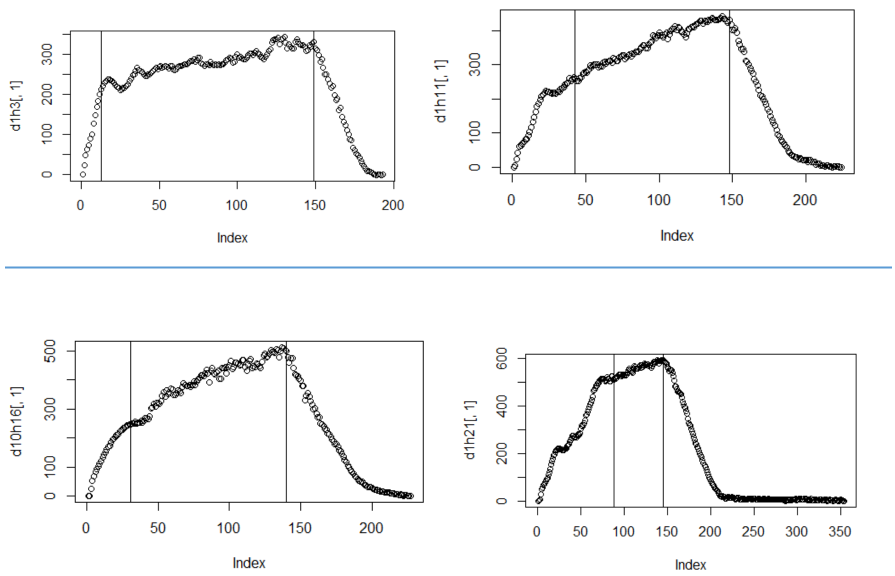


- We apply the same procedure to each dataset of each holes of all drill bits and find the middle stage data to be considered for modelling. The same threshold is taken for the graph of ‘torque’ versus ‘time elapsed’ as thrust force and torque have a good correlation in between them for all values of d[i]h[j]. This graph shows their correlation:



Formatiert: Schriftart: (Standard) Times New Roman, 12 Pt.

Some of the other examples of implementing the devised threshold for the taking out middle stage data are:



Thus, we get the desired data to be considered for model. What we do now is to take mean of those filtered data to get mean thrust force and mean torque for each set of d[i]h[j]. The final ready-to-use data table is as shown below:

	Drill.Number	Hole.present	Thrust.Mean	Torque.Mean	Row.Number	RUL	sqrt.thrust.mean	sqrt.torque.mean	percent
1	1	1	267.8195	16.25255	189	20	16.36519	4.031445	0.04761905
2	1	2	273.3961	18.55051	188	19	16.53469	4.307031	0.09523810
3	1	3	281.9069	18.61652	193	18	16.79008	4.314686	0.14285714
4	1	4	309.5438	21.84968	197	17	17.59386	4.674364	0.19047619
5	1	5	318.1396	22.61879	200	16	17.83647	4.755922	0.23809524
6	1	6	326.6118	23.23336	200	15	18.07240	4.820099	0.28571429
7	1	7	329.7669	24.14794	202	14	18.15949	4.914055	0.33333333
8	1	8	347.2958	24.40285	209	13	18.63587	4.939924	0.38095238
9	1	9	361.1649	25.73251	218	12	19.00434	5.072722	0.42857143
10	1	10	353.6191	25.12729	208	11	18.80476	5.012713	0.47619048
11	1	11	361.8815	25.20575	224	10	19.02318	5.020532	0.52380952
12	1	12	379.6081	26.72737	223	9	19.48353	5.169852	0.57142857
13	1	13	387.1257	25.89286	216	8	19.67551	5.088503	0.61904762
14	1	14	412.1486	29.63666	222	7	20.30144	5.443957	0.66666667
15	1	15	412.3361	29.12831	220	6	20.30606	5.397065	0.71428571
16	1	16	417.4533	30.27762	243	5	20.43167	5.502510	0.76190476

This data table contains the data for all holes of all drill sets as said earlier. The notation for data variables shown in this table can be understood using the explanation given below:

Drill.Number - Drill bit number
Hole.present - hole being drilled presently

Thrust.Mean - Mean of the extracted value of thrust

Torque.Mean - Mean of the extracted value of torque

Row.Number - number of observations for each hole

RUL - Remaining Useful Life of the drill bit

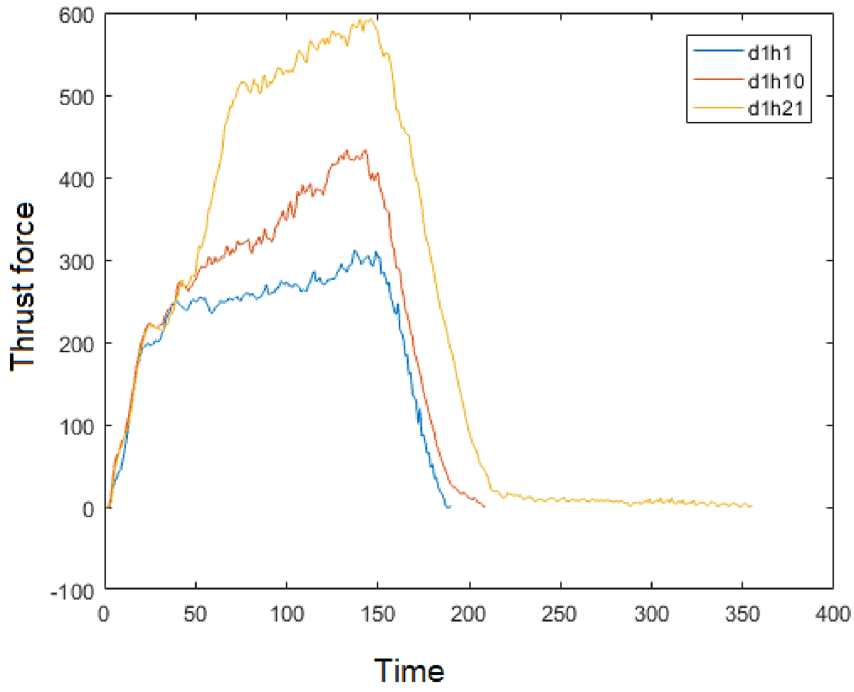
sqrt.thrust.mean = square root of "Thrust.Mean"

sqrt.torque.mean = square root of "Torque.Mean"

percent = percentage of life completed by the drill bit (Hole.present/(Hole.present + RUL)) [important]

WHY DID WE TAKE PERCENTAGE?

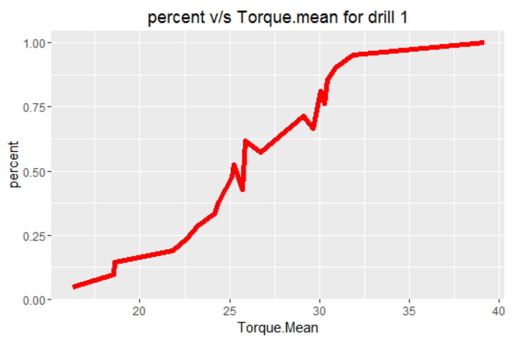
Let's take an example. When the number of drill bits increase, the corresponding maximum thrust force and torque increases. Let's take two datasets: d1h10 and d13h22. They have their current RUL as 11 and 23 respectively. We find that their respective mean torques are around 250 and 270 respectively. So, if we find RUL directly by training the model directly with thrust force and torque then we will be getting the predicted RUL values around 10 and 12 respectively. **Hence, we clearly see that finding RUL directly from data can give us wrong results. So, it's the percentage of journey the drill has completed which is directly correlated to the mean thrust force and mean torque.**

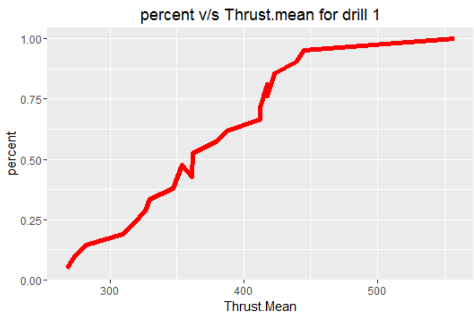


Formatiert: Schriftart: (Standard) Times New Roman, 12 Pt., Unterstrichen

WHY DID WE TAKE SQUARE ROOT IN MODEL

The graph of percent versus thrust.mean and torque.mean respectively are as follows which can be approximated as the graph of y = root (x).





5. MODELLING

Taking percent as dependent variable and thrust.mean, torque.mean, sqrt.thrust.mean, sqrt.torque.mean as independent variables, we make our model using linear regression as well as polynomial regression. The model on our training data set(drill 1 to drill 10) are as follows:

```
Call:
lm(formula = percent ~ Thrust.Mean + Torque.Mean + sqrt.thrust.mean +
    sqrt.torque.mean, data = train)

Residuals:
     Min       1Q   Median       3Q      Max
-0.42915 -0.09225 -0.00222  0.09021  0.68794

Coefficients:
                  Estimate Std. Error t value Pr(>|t|)
(Intercept)       0.245081   0.858532   0.285    0.776
Thrust.Mean       0.015357   0.003198   4.802 3.13e-06 ***
Torque.Mean      -0.131211   0.026068  -5.033 1.09e-06 ***
sqrt.thrust.mean -0.553476   0.128165  -4.318 2.50e-05 ***
sqrt.torque.mean  1.691827   0.285992   5.916 1.46e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.147 on 195 degrees of freedom
Multiple R-squared:  0.7467,    Adjusted R-squared:  0.7415
F-statistic: 143.7 on 4 and 195 DF,  p-value: < 2.2e-16
```

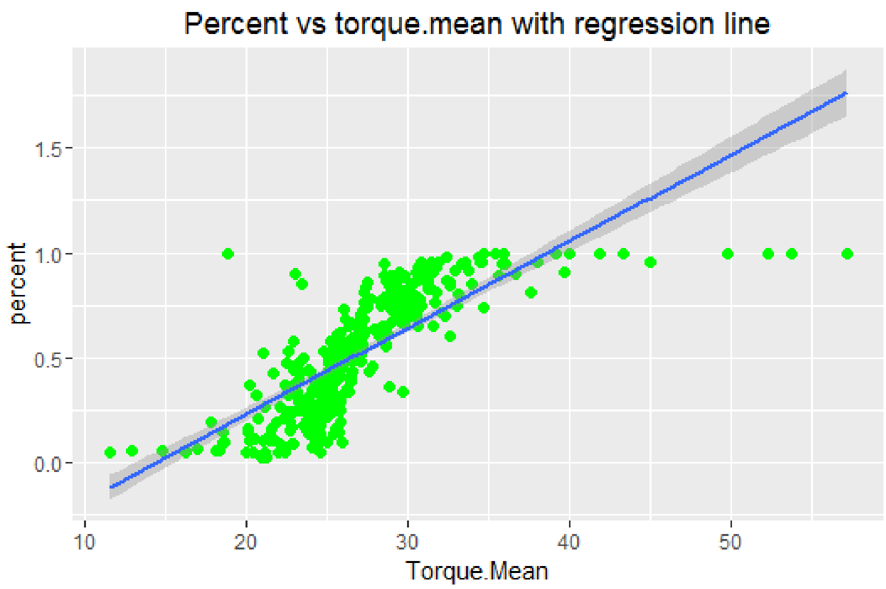
As we are applying polynomial regression, very few of the predicted RUL are coming less than zero and since these negative values have no physical significance, we are rounding off those values to zero. There are 4 such values to be precise.

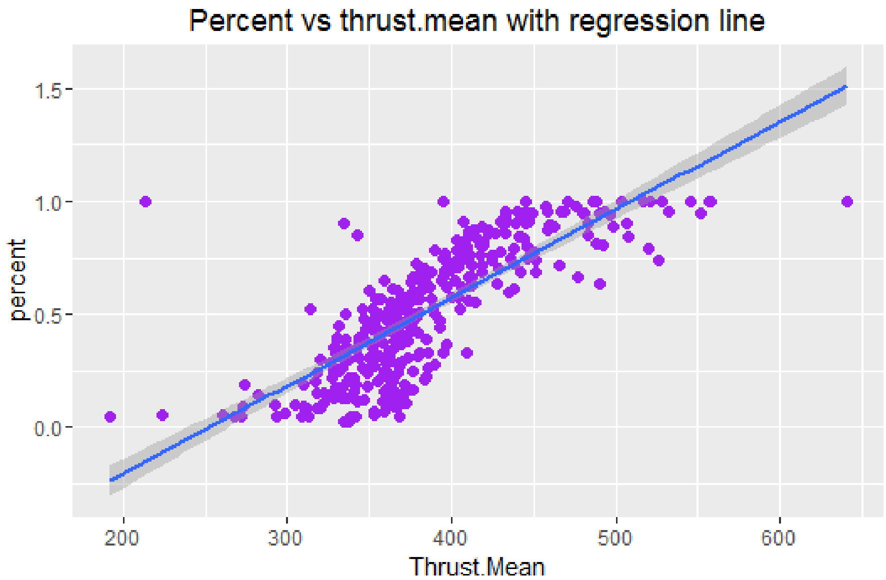
```
Console C:/Users/Neelam/Desktop/R directory/Drils and holes/last drills/
> table(RUL>0)

FALSE  TRUE
    4   129
>
```

Now, we compare some of the predicted values and actual values of RUL for drill 12.

Actual	Predicted
12	10
11	10
10	13
9	10
8	9
7	7
6	6
5	6

And again we compare the actual 'percent' variable and predicted 'percent' variable. The predicted one can be seen as a straight line as we have applied linear regression.





6. ANNEXURE

LINEAR REGRESSION: Linear regression is the most basic and commonly used predictive analysis. Regression estimates are used to describe data and to explain the relationship between one dependent variable and one or more independent variables.At the center of the regression analysis is the task of fitting a single line through a scatter plot. The simplest form with one dependent and one independent variable is defined by the formula

y = c + b*x, where y = estimated dependent, c = constant, b = regression coefficients, and x = independent variable.

Polynomial Regression: **polynomial regression** is a form of linear**regression** in which the relationship between the independent variable x and the dependent variable y is modelled as an nth degree **polynomial**.